Selection of Nontarget Arthropod Taxa for Field Research on Transgenic Insecticidal Crops: Using Empirical Data and Statistical Power

J. R. PRASIFKA,[1,2] R. L. HELLMICH,[1] G. P. DIVELY,[3] L. S. HIGGINS,[4] P. M. DIXON,[5] AND J. J. DUAN[6]

Environ. Entomol. 37(1): 1–10 (2008)

ABSTRACT One of the possible adverse effects of transgenic insecticidal crops is the unintended decline in the abundance of nontarget arthropods. Field trials designed to evaluate potential nontarget effects can be more complex than expected because decisions to conduct field trials and the selection of taxa to include are not always guided by the results of laboratory tests. Also, recent studies emphasize the potential for indirect effects (adverse impacts to nontarget arthropods without feeding directly on plant tissues), which are difficult to predict because of interactions among nontarget arthropods, target pests, and transgenic crops. As a consequence, field studies may attempt to monitor expansive lists of arthropod taxa, making the design of such broad studies more difficult and reducing the likelihood of detecting any negative effects that might be present. To improve the taxonomic focus and statistical rigor of future studies, existing field data and corresponding power analysis may provide useful guidance. Analysis of control data from several nontarget field trials using repeated-measures designs suggests that while detection of small effects may require considerable increases in replication, there are taxa from different ecological roles that are sampled effectively using standard methods. The use of statistical power to guide selection of taxa for nontarget trials reflects scientists' inability to predict the complex interactions among arthropod taxa, particularly when laboratory trials fail to provide guidance on which groups are more likely to be affected. However, scientists still may exercise judgment, including taxa that are not included in or supported by power analyses.

KEY WORDS experimental design, *Bacillus thuringiensis*, genetically modified, risk assessment

The potential for unintended declines in the abundance, activity, or diversity of arthropods is one of several issues considered in assessing the relative benefits and risks of transgenic (=genetically modified or genetically engineered) crop production (Wolfenbarger and Phifer 2000). Such adverse outcomes are generally referred to as nontarget (or nontarget) effects, which perhaps derives from the use of the term "nontarget organisms" to describe taxa impacted by the early, indiscriminate use of insecticides (Newsom 1967), and subsequently in importation (=classical) biological control (Howarth 1991, Louda et al. 2003). Because many of the commercially available transgenic crops express insecticidal proteins derived from the soil bacterium *Bacillus thuringiensis* (Berliner) (*Bt*), the term nontarget adequately describes all species that a novel genetic combination is not intended to suppress.

Using field tests to evaluate the potential effects of transgenic crops on nontarget arthropods might seem straightforward. Existing frameworks to assess risk have been adapted to nontarget taxa (Andow and Hilbeck 2004), and there is widespread agreement that such a process is best organized in a stepwise or tiered fashion (Romeis et al. 2006). In theory, field testing is conducted based on the need to clarify results from earlier, lower-tiered tests. Logically, the identity of species adversely affected in laboratory tests would help determine the list of taxa to include in more realistic field research. However, decisions to conduct field trials and the selection of taxa to include in such studies are often decoupled from the results of laboratory testing. Broader groups of arthropods are generally used to monitor for nontarget effects including taxa considered (1) likely to be exposed to an insecticidal toxin, (2) to provide ecosystem services, and (3) rare or charismatic. Although considering the likelihood of exposure and special human interests have merit, these criteria are often based on the opinions of experts, which may not be accurate or scientifically justifiable (as noted in Andow and Hilbeck 2004).

Furthermore, experience with *Bt* crops suggests that assessing unintended effects of transgenic insec-

Mention of a proprietary product does not constitute an endorsement or a recommendation by Iowa State University or USDA for its use.

[1] USDA–ARS, Corn Insects and Crop Genetics Research Unit, Genetics Laboratory c/o Insectary, Iowa State University, Ames, IA 50011.

[2] Corresponding author, e-mail: prasifka@iastate.edu.

[3] Department of Entomology, University of Maryland, College Park, MD 20742.

[4] Pioneer Hi-Bred International, Johnston, IA 50131.

[5] Department of Statistics, Iowa State University, Ames, IA 50011.

[6] Monsanto Company–Ecological Training Center, St. Louis, MO 63167.

ticidal crops in the field is more complex than anticipated and that selection of nontarget taxa is a critical step. For example, predictions regarding direct effects (i.e., those resulting from nontarget feeding on transgenic crop tissues) on species likely to be susceptible to a toxin might be misleading if field exposure (consumption of the toxin) is not accurately estimated (as in Losey et al. 1999, Jesse and Obrycki 2000; see Hellmich et al. 2001). Among indirect effects, there is potential for reduced populations of predators or parasitoids that depend on target pests as prey or hosts (Riddick et al. 1998, Wold et al. 2001, Pilcher et al. 2005). The conclusion that nontarget arthropods may be adversely affected through consuming prey containing *Bt* toxins (Hilbeck et al. 1998, Dutton et al. 2002, Ponsard et al. 2002) is potentially more problematic. Although none of these studies proves any environmental harm from *Bt* crops and similar effects on nontarget taxa might be expected from other pest management strategies, the scientific discussion has shifted to the potential for indirect effects. This focus on indirect effects is perhaps responsible for the cautious, more expansive lists of arthropod taxa monitored (see Dively 2005).

Evaluation of how nontarget effects are assessed should ideally be objective, but disadvantages of de facto monitoring of all identifiable taxa exist for both supporters and detractors of transgenic crops. Drawbacks for biotechnology advocates, particularly industry groups, include a lack of guidance in designing such broad studies and the taxonomic expertise necessary to conduct them. For opponents of transgenic crops, one major shortcoming is that increasingly broad efforts may be unable to detect any negative effects that might be present. The demand to test for treatment effects on more taxa also may result in weak analysis of rare or poorly sampled taxa. In some cases, failure to detect such effects may incorrectly be equated with an absence of effects (Marvier 2002), but each distinct test also presents an additional chance for a type I error (incorrect rejection of a true null hypothesis; i.e., apparent detection of an effect that does not exist).

If field testing continues to be used to search for indirect nontarget effects of transgenic insecticidal crops, changes to the current research methods should be considered. In particular, studies lacking well-defined hypotheses (including broad monitoring efforts that may be almost limitless in scope) could benefit by applying more objective grounds for selecting nontarget taxa and focusing on fewer total nontarget groups. The quality with which various arthropods are sampled is one quantitative criterion for selecting nontarget taxa. Statistical power, which represents the probability that an incorrect null hypothesis (e.g., that nontarget densities among treatments are similar) will be correctly rejected by a particular test, can indicate the quality of sampling in a way that addresses the adequacy of experimental designs. Using existing data sets, power analysis may be used to help select a modest number of nontarget taxa and improve the soundness of hypothesis testing. An example using this concept with historical sampling data from nontarget studies on transgenic field corn is described below.

Materials and Methods

A total of 15 time series (5 locations × 3 yr) data sets containing abundance estimates for arthropod taxa were collected. In an effort to make more generally applicable conclusions, study locations included several Corn Belt states (Iowa, Nebraska, Illinois, Maryland) and types of research groups (federal, state, industry). Although some of the data come from previously published research, the inclusion of sampling data used to evaluate proprietary transgenic varieties dictated that only data from negative control treatments (i.e., no insecticidal transgene, no conventional insecticides) were included. A summary of the nontarget sampling protocols for the included data sets is provided in Table 1. As a preliminary step in summarizing the data sets, simple statistics (mean, CV) were calculated for all taxa and sampling date combinations. Some studies included abundance estimates using uncommon or novel sampling techniques, but only abundance estimates using the most common methods (pitfall traps, sticky cards, visual counts) were evaluated.

Because some data sets provided abundance estimates for 100 or more nontarget groups (with some taxa sampled using multiple methods), a need to initially narrow the list of candidate nontarget groups was apparent. Although other nonmutually exclusive methods could be used, in this case taxon × method combinations were excluded that did not meet the minimum criteria of being sampled (1) at two or more locations, (2) with an observed CV <100 for at least two consecutive sampling periods within a season. These conditions aimed to eliminate taxa that were not common over a broad geographic range or sampled with a minimum level of precision. One additional criterion rejected taxon × method combinations for which (3) the condition that CV < 100 (over consecutive samples) was true in less than two thirds of the location × year combinations. This restriction helped exclude for which the quality of sampling was not consistent from year-to-year. Although the criteria used are somewhat arbitrary, they agree in concept with the association of taxa with CV < 100 having high statistical power (Duan et al. 2006) and reduced the number of candidate by ≈80%.

The resulting list of nontarget taxa, separated by ecological role (herbivores, saprovores, predators, parasitoids; equivalent to "functional group" in Dively 2005), includes several taxa that did not meet the minimum criteria but were considered to be of special interest (Table 2). Many of the taxa are grouped at the family level, which could be criticized as too broad. However, resolution of data to family was sometimes necessary to compare among data sets when common genera or species differed among locations and may be appropriate when previous testing does not indicate which arthropods are most likely to be adversely affected in field tests. Also, though classification of some taxa (e.g., carabid beetles) using a single ecological

Table 1. Summary of nontarget sampling protocols at five study locations, 2000–2003

Location	Replicates	Plot size (ha)	Sampling methods	Sample periods[a] (year 1, 2, 3)	Subsamples per plot (duration)			
					2000	2001	2002	2003
Illinois	4	0.34 ha	Sticky cards	4, 5, 4	3 (7 d)	3 (7 d)	3 (7 d)	
			Pitfall traps	4, 4, 4	4 (3 d)	4 (3 d)	4 (3 d)	
			Root/soil samples	3, 3, 3	3	3	3	
Iowa-1	3–4	0.13–0.37 ha	10-plant visual	17, 16, 6		1	1	11
			Sticky cards	10, 7, 6		10 (7 d)	10 (7 d)	11 (1 d)
			Pitfall traps	10, 8, 6		10 (7 d)	10 (7 d)	11 (1 d)
			Straw litter bags	1				4 (38–49 d)[c]
Iowa-2	2	0.03–0.04 ha	20-plant visual	7, 7, 4		4	4	4[b]
			Sticky cards	5, 5, 4		4 (1 d)	4 (1 d)	10 (1 d)
			Pitfall traps	10, 10, 4		4 (1 d)	4 (1 d)	10 (1 d)
			Soil samples	10, 10		4	4	
			Straw litter bags	4				4 (31–75 d)[c]
Maryland	3	0.38 ha	10-plant visual	9, 12, 4	1	1[b]	1[b]	
			Sticky cards	9, 12, 4	6 (7 d)	8 (7 d)	8 (7 d)	
			Pitfall traps	8, 11, 6	10 (7 d)	8 (7 d)	8 (7 d)	
			Litter samples	1, 4, 1	8	8	4	
Nebraska	2	0.03–0.04 ha	20-plant visual	7, 7, 4		4	4	4[b]
			Sticky cards	5, 5, 4		4 (1 d)	4 (1 d)	10 (1 d)
			Pitfall traps	10, 10, 4		4 (1 d)	4 (1 d)	10 (1 d)
			Soil samples	10, 10		4	4	
			Straw litter bags	4				4 (32–91 d)[c]

[a] Number of sample periods for each sampling method.
[b] Visual counts used 10-plants at Nebraska and Iowa (1) in 2003 and 8-plants at Maryland in 2001–2002.
[c] Straw litter bags placed both above and below ground and remained in plots for a range of days.

role can seem inaccurate or misleading, this was a necessary simplification; for many taxa, genus or species level identifications could not be made without additional taxonomic expertise, meaning more precise assignment of ecologic roles was not possible. Furthermore, the use of simplified ecological roles should prove useful. By including members with differing ecological roles, taxonomic, and ecological breadth is likely even for short lists of nontarget arthropods. Given differences in life histories and the distribution of resources among taxa with distinct ecological roles (Price 1976), sampling distributions and power for a

Table 2. Nontarget arthropod taxa included for prospective power analyses

Ecological role[a]	Taxon	Common name	Sampling methods (life stages)[b]
Herbivores	Heteroptera: Cicadellidae	Leafhoppers	Sticky cards (i a)
	Heteroptera: Aphididae	Aphids	Sticky cards (i a)
	Heteroptera: Fulgoroidea	Planthoppers	Sticky cards (i a)
	Thysanoptera: Thripidae	Thrips	Sticky cards (i a)
	Coleoptera: Chrysomelidae: *Diabrotica*	Corn rootworms	Visual counts, sticky cards
	Coleoptera: Chrysomelidae: Alticinae	Flea beetles	Visual counts, sticky cards[c]
Saprovores	Arachnida: Opiliones	Harvestmen	Pitfall traps[c]
	Collembola	Springtails	Pitfall traps (i a)
	Collembola: Entomobryidae	Entomobryids	Pitfall traps (i a)
	Coleoptera: Nitidulidae	Sap beetles	Pitfall traps[c]
	Orthoptera: Gryllidae	Crickets	Pitfall traps (i a)
	Diptera: Chloropidae	Frit flies	Sticky cards
	Diptera: Sciaridae	Fungus gnats	Sticky cards
	Hymenoptera: Formicidae	Ants	Pitfall traps
Predators	Arachnida: Araneae	Spiders	Pitfall traps (i) (a), visual counts
	Arachnida: Araneae: Lycosidae	Wolf spiders	Pitfall traps[c]
	Chilopoda	Centipedes	Pitfall traps[c]
	Neuroptera: Chrysopidae	Green lacewings	Visual counts (e)[c]
	Heteroptera: Anthocoridae: *Orius*	Minute pirate bugs	Sticky cards
	Coleoptera: Carabidae	Ground beetles	Pitfall traps (i)[c] (a)
	Coleoptera: Staphylinidae	Rove beetles	Pitfall traps
	Coleoptera: Coccinellidae	Ladybird beetles	Visual counts (i) (a), sticky cards
	Diptera: Dolochopodidae	Long-legged flies	Sticky cards
Parasitoids	Hymenoptera: Mymaridae	Fairyflies	Sticky cards
	Hymenoptera: Scelionidae	Scelionids	Sticky cards
	Hymenoptera: Braconidae	Braconids	Sticky cards[c]
	Hymenoptera: Trichogrammatidae	Trichogrammatids	Sticky cards[c]

[a] Assignment to a primary group was made for taxa with a variety of ecological roles (e.g., carabid beetles).
[b] Adults unless otherwise indicated. Life stages (eggs, immatures, adults) combined if paired inside parentheses.
[c] Did not meet minimum criteria for inclusion but added as group or stage of special interest.

given experimental design also are likely to differ (e.g., between herbivores and parasitoids; Gould and Naranjo 1999).

Power Analyses. Prospective analyses using the PASS software package (NCSS 2002) were used to estimate the power of hypothesis tests in similar nontarget studies that might be conducted in the future. The power analyses assumed treatment effects on individual taxa would be evaluated using a repeated-measures analysis (RM-analysis of variance [ANOVA]) to test for a difference between the negative control and one experimental treatment (a transgenic insecticidal cultivar). To include a range of possible outcomes, analyses estimated power if the transgenic variety reduced the overall mean for a nontarget taxon by 20, 30, or 50%, including 3, 6, 9, 12, and 15 replicates of each treatment. Although β (the type II error rate) varied to defined power as $1 - \beta$, α (the type I error rate) was set at 0.05 for all analyses. To produce the desired output (plots of power versus replication) PASS further required (1) the between-subjects mean square (MSB), a measure of variation among replicates (NCSS 2002), (2) the number of time periods or repeated measurements, and (3) the treatment means being compared. For a single experiment and nontarget group, this is relatively simple with MSB (estimated by the mean square of the replicate or block $\times$ treatment effect) and mean selected from a RM-ANOVA output.

However, because >300 specific combinations of taxon, location, year, and sampling method were analyzed, an SAS program (SAS Institute 1999; PROC GLM) was used to generate only the information necessary as inputs for the power analyses. For all of the location $\times$ year combinations containing observations for each taxon (and sampling method), the original abundance estimates were used to create four variables. The first variable, y_1, was the log-transformed control data, with y_2– y_4 representing treatments with 20, 30, or 50% reductions in nontarget arthropod abundance (percentages before transformation; i.e., $y_1 = \log [x + 1]$, $y_2 = \log [0.8x + 1]$, $y_3 = \log [0.7x + 1]$, $y_4 = \log [0.5x + 1]$). The logarithmic transformation was applied to abundance data to compensate for the frequent problems of right-skewed distributions or positive correlations between means and variances (see Sokal and Rohlf 1995). For each y_1–y_4, the mean and the between-subjects mean square were output. To obtain the relationship between power and replication for a −20% effect, inputs into PASS included (1) the means y_1 and y_2, (2) the number of sampling periods, and (3) the arithmetic mean of MSB for y_1 and y_2. For other effect sizes, the corresponding pairs of transformed means (i.e., y_1 and y_3 for a −30% effect, y_1 and y_4 for a −50% effect) were used to estimate the effect of replication on statistical power.

Outputs and Assumptions. For each nontarget taxon and sampling method, PASS outputs were comprised of several ($n = 6$–15) estimates of the relationship between power and replication. Because each of the estimates corresponded to a distinct data set, variation among them could come from several sources, including year-to-year differences in the abundance and distribution of nontarget arthropod populations or greater sampling effort at specific locations. Consequently, the estimate with the median power to detect reductions in abundance for a nontarget group (across all combinations of effect and replication) was selected to represent the relationship between power and replication for that taxon. Plots of power versus replication (power curves) were used to graphically approximate the power of hypothesis tests in similar future nontarget studies. The largest potential change in abundance of nontarget taxa (−50%) is representative of a direct effect, but power curves intended to represent indirect effects (−20, −30%) of transgenic crops were also generated. The median curve was preferred because it retains a more realistic shape of the relationship between power and replication compared with a curve based on the mean power at each level of replication (which would appear relatively flattened).

Several specific assumptions are required for repeated-measures experiments and the corresponding power analyses (NCSS 2002). Although violations of some assumptions may produce only minor changes to results, these (and other) power analyses should be considered optimistic or "best-case" assessments of statistical power. For example, monitoring for some taxa produced abundance estimates of zero for specific dates, particularly in early samples; because the PASS inputs included the number of sample dates and overall (seasonal) treatment means, in such cases the output overestimates the probability of detecting a treatment difference. Rather than apply subjective judgment across nontarget groups in all of the data sets by deleting or otherwise modifying specific observations, each data set was accepted without any changes. Even with the potential for violating one or more assumptions implicit in this approach, the overall quality and quantity of information should provide results useful for selecting a limited number of nontarget taxa for future field research.

Results and Discussion

To place prospective power estimates into context, statistical power of at least 0.70 has been suggested for field research on nontarget effects (Perry et al. 2003, Duan et al. 2006). Aside from the level of variability within the data (represented by MSB in the PASS analyses), power largely depends on the magnitude of the difference between treatments; unfortunately, the effect size considered probable or biologically significant is generally unclear (Perry et al. 2003). Consequently, defaulting to an effect size (−50 or −30%; Lopez et al. 2005, Duan et al. 2006) for which high ($\geq$0.70) power can be achieved seems likely unless more biologically meaningful guidelines are developed.

Recommended Taxa and Sampling Methods. Power curves were generated for the three taxa with the greatest median power to detect decreases in abun-

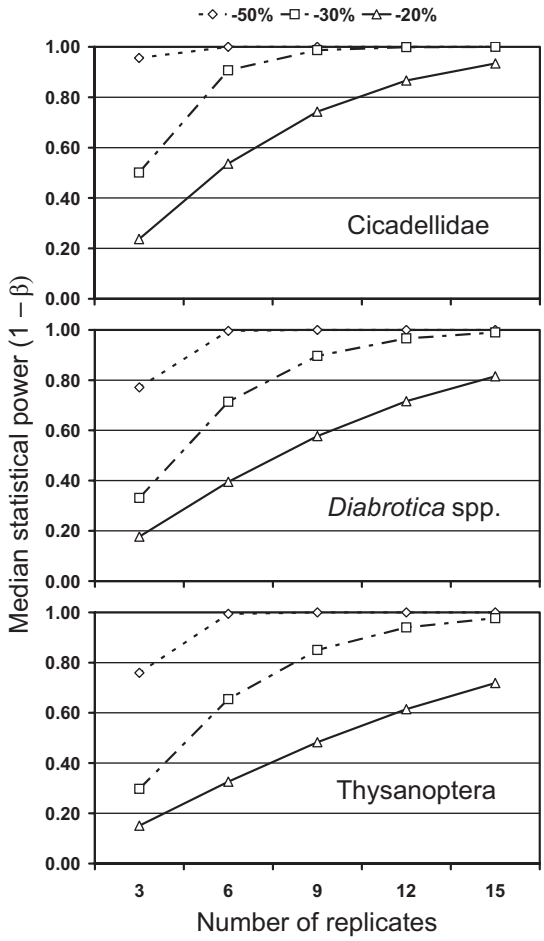

Fig. 1. Relationship between power (1 − β) and replication for nontarget herbivores. Plotted curves estimate the median power to detect a −50, −30, or −20% change in abundance for indicated nontarget taxa sampled using sticky cards.

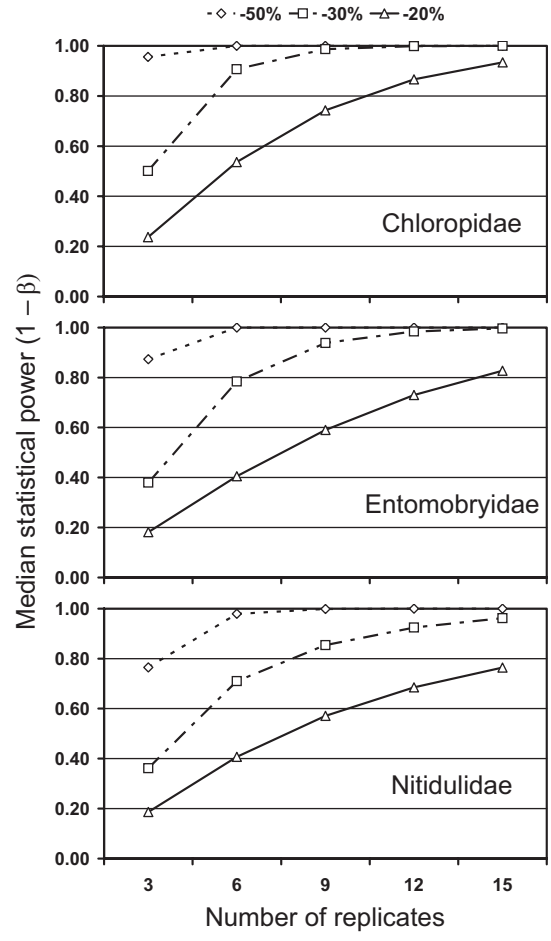

Fig. 2. Relationship between power (1 − β) and replication for nontarget saprovores. Plotted curves estimate the median power to detect a −50, −30, or −20% change in abundance for indicated nontarget taxa sampled using sticky cards (Chloropidae) or pitfall traps (Collembola, Nitidulidae).

dance for each of the four ecological roles. Estimates of the median power versus replication relationships for all taxa tested can be found in Appendix 1.

Among the herbivores, leafhoppers (Cicadellidae), corn rootworm adults (*Diabrotica* spp.), and common thrips (Thripidae) seem to be sampled most effectively (Fig. 1), all using sticky cards to estimate abundance. Because *Diabrotica* spp. is a target pest for some *Bt* corn varieties, flea beetles (Chrysomelidae: Alticinae; sampled with sticky cards) may be included as another nontarget herbivore. For all four herbivores, 80% power was estimated for detecting a large (−50%) effect with three to four replicates, whereas approximately nine replicates would be needed to detect more modest (−30%) changes. For a given level of replication, visual counts for corn rootworm adults and flea beetles seemed less effective than traps and probably required more in-field effort.

The prospective power estimates for grass flies (Chloropidae) sampled with sticky cards and springtails (Collembola) and sap beetles (Nitidulidae) sampled with pitfall traps were highest among decomposers (Fig. 2). Pitfall sampling for a single family of springtails (Entomobryidae) also appeared effective, with power >0.70 to detect a 30% decrease with approximately six replicates. If *Bt* corn varieties targeting the European corn borer (*Ostrinia nubilalis* Hübner) reduce damage to corn ears, sap beetle abundance may be reduced (Daly and Buntin 2005). In such a case, fungus gnats (Sciaridae) could be a more informative nontarget saprovore, although at similar effect sizes greater replication seems to be needed. However, some fungus gnats may be indirectly affected by the use of *Bt* crops (Mycetophilidae; Candolfi et al. 2004).

Minute pirate bugs (*Orius* spp., sticky cards), ladybird beetle adults (Coccinellidae, visual counts), and wolf spiders (Lycosidae, pitfall traps) were the most effectively sampled predators, using different sam-

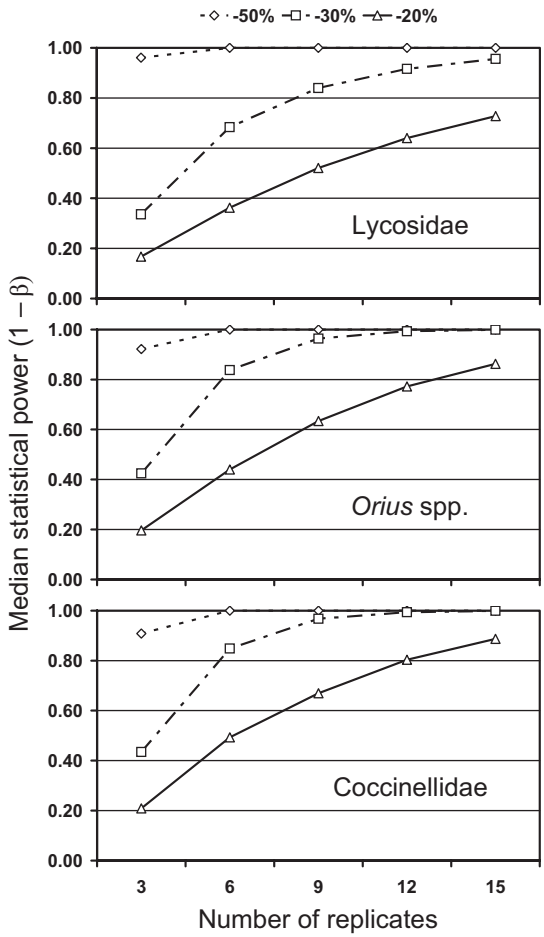

Fig. 3. Relationship between power $(1 - \beta)$ and replication for nontarget predators. Plotted curves estimate the median power to detect a −50, −30, or −20% change in abundance for indicated nontarget taxa sampled using sticky cards (*Orius* spp.), visual counts (Coccinellidae), or pitfall traps (Lycosidae).

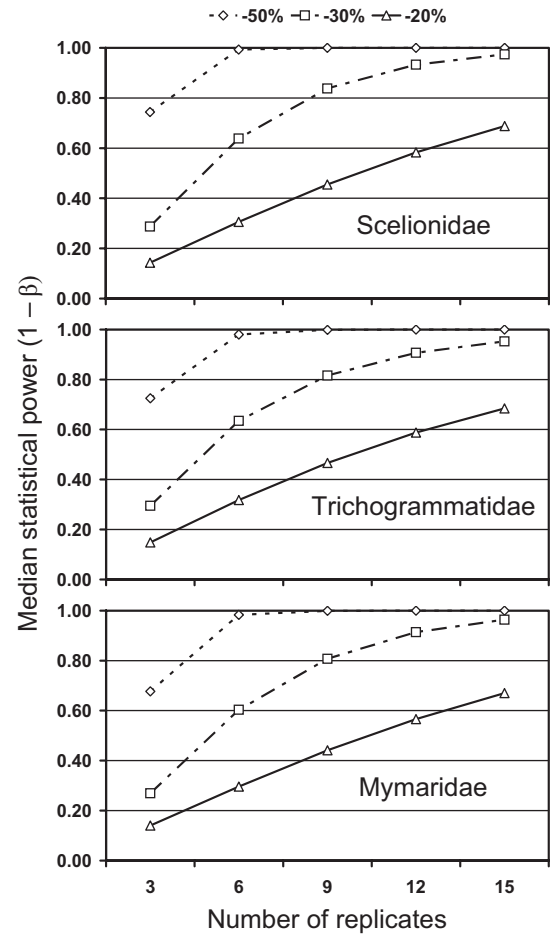

Fig. 4. Relationship between power $(1 - \beta)$ and replication for nontarget parasitoids. Plotted curves estimate the median power to detect a −50, −30, or −20% change in abundance for indicated nontarget taxa sampled using sticky cards.

pling methods suggested for each group (Fig. 3). Power analyses of all spiders together provided results similar to those for wolf spiders, likely because lycosids comprised the largest single group of spiders in several data sets. Pitfall trap results also suggested ground (carabid) and rove (staphylinid) beetles required an estimated four or six replicates, respectively, to detect a −50% change in abundance with over 70% probability $(1 - \beta \geq 0.70)$.

At a given effect size, estimates of power for parasitoids in the families Scelionidae, Trichogrammatidae, and Mymaridae (fairyflies) were generally lower or increased with replication more slowly (Fig. 4) than those for the representatives with other ecological roles (Figs. 1–3). Sampling for braconids seemed less efficient than the other parasitoids, all of which were evaluated using data from sticky cards. Data on braconid abundance were included for 12 of the 15 data sets, but in some cases, only a single species (*Macrocentrus cingulum* Reinhard) was counted, perhaps contributing to lower predicted power relative to the other three parasitoid families.

Application of Results to Nontarget Studies. Several conclusions are supported by the results of the prospective power analyses. First, it seems some taxa from all of the ecological roles can be sampled with adequate power to detect large (−50%) changes with only three to four replicates. However, in most cases, 10 or more replicates will be required to detect small (20%) reductions in the abundance of nontarget taxa. Also, some taxa initially excluded for failure to meet minimum sampling criteria (sap beetles, wolf spiders) showed particularly high predicted power.

In general, the groups that showed the best predicted power versus replication relationships are likely to be sampled effectively in future studies if comparable methods (Table 1) are used. However, because of differences among the included studies (location, sampling methods), other taxa not tested or

not supported by the power analyses may be useful as indicators of possible nontarget effects. Also, although more specialized methods including soil cores, litter samples, and straw litter bags were not used at enough locations to be included in the overall power analysis, other results suggest these methods may be equivalent or better sampling methods for certain nontarget taxa (Prasifka et al. 2007).

Statistical Power and Alternate Analyses. Other studies that have used power analysis support the conclusion that increased replication is necessary to detect small or moderate effects on nontarget arthropod abundance (Bourguet et al. 2002, Perry et al. 2003, Lopez et al. 2005, Naranjo 2005). Perry et al. (2003) used a simulation approach that tested power using several models including various combinations of mean, variance, and effect size to help determine appropriate replication needs for field-scale nontarget trials. Including field abundance data from many different taxonomic groups, Duan et al. (2006) focused on retrospectively assessing the power of a nontarget study. Analysis of field data by Naranjo (2005) showed relatively greater gains in statistical power by increasing replication rather than increasing the number of sample dates (repeated measures). The simplified approach here uses a tool for power analysis accessible to scientists without statistical specialization (NCSS 2002), but may provide more broadly applicable results by incorporating field data that encompassed variation in time, space, and methods. Without regard to the specific approach used to estimate statistical power, using empirical data to help select nontarget taxa for field studies can aid in designing future field experiments, because choices of nontarget taxa to include in a study necessarily impact other aspects of experimental design (Andow and Hilbeck 2004, Prasifka et al. 2005). Criticism of experimental design elements as inadequate or inappropriate (duration, plot size, sampling methods; EPA 2001, 2002) also suggest such an integrated approach would be beneficial.

Although much attention has been given to the relationship between statistical power and replication, other approaches to improve power should be acknowledged. As an alternative to separate analyses of study years, Duan et al. (2006) estimated pooling 2 yr of a nontarget study increased the percentage of hypothesis tests with satisfactory power from 22 to 86% (also see Naranjo 2005). Meta-analysis (Hunter et al. 1982, Hedges and Olkin 1985, Marvier et al. 2007) can increase statistical power by integrating experimental results across similar nontarget studies. Multivariate procedures, such as principal response curve analysis (PRC; Van den Brink and ter Braak 1999) assess nontarget impacts at the community level (Naranjo et al. 2003, Dively 2005, Naranjo 2005, Prasifka et al. 2005, Torres and Ruberson 2005, Whitehouse et al. 2005). Using PRC may be more powerful when several sampled taxa respond to an experimental treatment, but this approach may be less powerful if one or very few taxa are impacted (Ammann et al. 2001).

An emphasis on using empirical data and statistical power to guide nontarget trials, particularly for selection of included taxa, might seem (incorrectly) to suggest the removal of biological expertise and reasoning from research with transgenic insecticidal crops. More accurately, the use of statistical power to guide selection of taxa for nontarget trials reflects scientists' inability to predict complex interactions among nontarget groups, target pests, and transgenic crops. It is intended to initiate selection of taxa for studies using in-field monitoring, particularly when earlier, lower-tiered testing fails to indicate which nontarget groups are most likely to be affected. Such a method does not preclude using judgment; additional taxa that are difficult to sample effectively (or for which relative sampling efficiency is not known) may subsequently be included because of their perceived value to humans. Prior experience with transgenic crops suggests such a balanced approach is essential to successful resolution of an issue of scientific and public interest (Pew Initiative on Food and Biotechnology 2002).

Acknowledgments

This research was a joint contribution from the USDA Agricultural Research Service and the Iowa Agriculture and Home Economics Experiment Station, Ames, IA (Project 3543). This article reports the results of research only. Valuable assistance in simplifying the SAS analysis was provided by D. Sumerford. Research was funded by the USDA Biotechnology Risk Assessment and Research Grants Program Grant 2004-39454-15193.

References Cited

Ammann, L. P., J. Kennedy, and W. T. Waller. 2001. Power comparison of principal response curves and ANOVA methods (http://www.utdallas.edu/~amman/Rabundance/Setac2001).

Andow, D. A., and A. Hilbeck. 2004. Science-based risk assessment for nontarget effects of transgenic crops. Bioscience 54: 637–649.

Bourguet, D., J. Chaufaux, A. Micoud, M. Delos, B. Naibo, F. Bombarde, G. Marque, N. Eychenne, and C. Pagliari. 2002. *Ostrinia nubilalis* parasitism and the field abundance of non-target insects in transgenic *Bacillus thuringiensis* corn (*Zea mays*). Environ. Biosafety Res. 1: 49–60.

Candolfi, M. P., K. Brown, C. Grimm, B. Reber, and H. Schmidli. 2004. A faunistic approach to assess potential side-effects of genetically modified Bt-corn on non-target arthropods under field conditions. Biocontrol Sci. Technol. 14: 129–170.

Daly, T., and G. D. Buntin. 2005. Effect of *Bacillus thuringiensis* transgenic corn for lepidopteran control on nontarget arthropods. Environ. Entomol. 34: 1292–1301.

Dively, G. P. 2005. Impact of VIP3A × Cry1Ab lepidopteran-resistant field corn on the nontarget arthropod community. Environ. Entomol. 34: 1267–1291.

Duan, J. J., C. Jiang, G. P. Head, M. A. Bhatti, D. P. Ward, S. L. Levine, T. E. Nickson, and M. A. Nemeth. 2006. Statistical power analysis of a two-year field study and design of experiments to evaluate non-target effects of genetically modified *Bacillus thuringiensis* corn. Ecol. Entomol. 31: 521–531.

Dutton, A., H. Klein, J. Romeis, and F. Bigler. 2002. Uptake of Bt-toxin by herbivores feeding on transgenic maize and consequences for the predator *Chrysoperla carnea.* Ecol Entomol. 27: 441–447.

Environmental Protection Agency [EPA]. 2001. FIFRA Scientific Advisory Panel Meeting Minutes, October 18, 2000, *Bt* plant-pesticides risk and benefit assessments. EPA, Arlington, VA.

Environmental Protection Agency [EPA]. 2002. FIFRA Scientific Advisory Panel Meeting Minutes, August 27, 2002, Part A: corn rootworm plant-incorporated protectant non-target insect and insect resistance management issues: non-target insect issues. EPA, Arlington, VA.

Gould, J. R., and S. E. Naranjo. 1999. Distribution and sampling of *Bemisia argentifolii* (Homoptera: Aleyrodidae) and *Eretmocerus eremicus* (Hymenoptera: Aphelinidae) on cantaloupe vines. J. Econ. Entomol. 92: 402–408.

Hedges, L., and I. Olkin. 1985. Statistical methods for meta-analysis. Academic, New York.

Hellmich, R. L., B. D. Siegfried, M. K. Sears, D. E. Stanley-Horn, M. J. Daniels, H. R. Mattila, T. Spencer, K. G. Bidne, and L. C. Lewis. 2001. Monarch larvae sensitivity to *Bacillus thuringiensis*-purified proteins and pollen. Proc. Natl. Acad. Sci. U.S.A. 98: 11925–11930.

Hilbeck, A., M. Baumgartner, P. Fried, and F. Bigler. 1998. Effects of transgenic *Bacillus thuringiensis* corn-fed prey on mortality and development time of immature *Chrysoperla carnea* (Neuroptera: Chrysopidae). Environ. Entomol. 27: 480–487.

Howarth, F. G. 1991. Environmental impacts of classical biological control. Annu. Rev. Entomol. 36: 485–509.

Hunter, J., F. Schmidt., and G. Jackson. 1982. Meta-analysis: cumulating research findings across studies. Sage Publications, Beverly Hills, CA.

Jesse, L., and J. Obrycki. 2000. Field deposition of Bt transgenic corn pollen: lethal effects on the monarch butterfly. Oecologia (Berl.) 125: 241–248.

Lopez, M. D., J. R. Prasifka, D. J. Bruck, and L. C. Lewis. 2005. Utility of ground beetle species as indicators of potential non-target effects of Bt crops. Environ. Entomol. 34: 1317–1324.

Losey, J., L. Rayor, and M. Carter. 1999. Transgenic pollen harms monarch larvae. Nature (Lond.) 399: 214.

Louda, S. M., R. W. Pemberton, M. T. Johnson, and P. A. Follett. 2003. Nontarget effects—the Achilles' heel of biological control? Retrospective analyses to reduce risk associated with biocontrol introductions. Annu. Rev. Entomol. 48: 365–396.

Marvier, M. 2002. Improving risk assessment for nontarget safety of transgenic crops. Ecol. Appl. 12: 1119–1124.

Marvier, M., C. McCreedy, J. Regetz, and P. Kareiva. 2007. A meta-analysis of effects of Bt cotton and maize on nontarget invertebrates. Science 316: 1475–1477.

Naranjo, S. E. 2005. Long-term assessment of the effects of transgenic *Bt* cotton on the abundance of non-target arthropod natural enemies. Environ. Entomol. 34: 1193–1210.

Naranjo, S. E., J. R. Hagler, and P. C. Ellsworth. 2003. Improved conservation of natural enemies with selective management systems for *Bemisia tabaci* (Homoptera: Aleyrodidae) in cotton. Biocontrol Sci. Technol. 13: 571–587.

NCSS. 2002. PASS 2002: power analysis and sample size for Windows user's guide—II. NCSS, Kaysville, UT.

Newsom, L. D. 1967. Consequences of insecticide use on nontarget organisms. Annu. Rev. Entomol. 12: 257–286.

Perry, J. N., P. Rothery, S. J. Clark, M. S. Heard, and C. Hawes. 2003. Design, analysis and statistical power of the farm-scale evaluations of genetically modified herbicide-tolerant crops. J. Appl. Ecol. 40: 17–31.

Pew Initiative on Food and Biotechnology. 2002. Three years later: genetically engineered corn and the Monarch butterfly controversy. Pew Initiative on Food and Biotechnology, Washington, DC.

Pilcher, C. D., M. E. Rice, and J. J. Obrycki. 2005. Impact of transgenic *Bacillus thuringiensis* corn and crop phenology on five nontarget arthropods. Environ. Entomol. 34: 1302–1316.

Ponsard, S., A. Gutierrez, and N. Mills. 2002. Effect of Bt-toxin (Cry1Ac) in transgenic cotton on the adult longevity of four heteropteran predators. Environ. Entomol. 31: 1197–1205.

Prasifka, J. R., R. L. Hellmich, G. P. Dively, and L. C. Lewis. 2005. Assessing the effects of pest management on non-target arthropods: the influence of plot size and isolation. Environ. Entomol. 34: 1181–1192.

Prasifka, J. R., M. D. Lopez, R. L. Hellmich, L. C. Lewis, and G. P. Dively. 2007. Comparison of pitfall traps and litter bags for sampling ground-dwelling arthropods. J. Appl. Entomol. 131: 115–120.

Price, P. W. 1976. Colonization of crops by arthropods: non-equilibrium communities in soybean fields. Environ. Entomol. 5: 605–611.

Riddick, E. W., G. Dively, and P. Barbosa. 1998. Effect of a seed-mix deployment of Cry3A-transgenic and non-transgenic potato on the abundance of *Lebia grandis* (Coleoptera: Carabidae) and *Coleomegilla maculata* (Coleoptera: Coccinellidae). Ann. Entomol. Soc. Am. 91: 647–653.

Romeis, J., M. Meissle, and F. Bigler. 2006. Transgenic crops expressing *Bacillus thuringiensis* toxins and biological control. Nat. Biotechnol. 24: 63–71.

SAS Institute 1999. SAS OnlineDoc, Version 8. SAS Institute, Cary, NC.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research, 3rd ed. W. H. Freeman and Company, New York.

Torres, J. B., and J. R. Ruberson. 2005. Canopy- and ground-dwelling predatory arthropods in commercial *Bt* and non-*Bt* cotton fields: patterns and mechanisms. Environ. Entomol. 34: 1242–1256.

Van den Brink, P. J., and C.J.F. ter Braak. 1999. Principal response curves: analysis of time-dependent multivariate responses of biological community to stress. Environ. Toxicol. Chem. 18: 138–148.

Whitehouse, M.E.A., L. J. Wilson, and G. P. Fitt. 2005. A comparison of arthropod communities in transgenic Bt and conventional cotton in Australia. Environ. Entomol. 34: 1224–1241.

Wold, S. J., E. C. Burkness, W. D. Hutchison, and R. C. Venette. 2001. In-field monitoring of beneficial insect populations in transgenic corn expressing a *Bacillus thuringiensis* toxin. J. Entomol. Sci. 36: 177–187.

Wolfenbarger, L. L., and P. R. Phifer. 2000. The ecological risks and benefits of genetically engineered plants. Science 290: 2088–2093.

Received for publication 27 February 2007; accepted 17 October 2007.

Appendix 1. Estimates of median statistical power versus replication for nontarget taxa listed in Table 2

Common name	Sampling method (life stages)[a]	Effect size[b]	Number of replicates				
			3	6	9	12	15
Leafhoppers	Sticky traps (i a)	−50%	0.956	1.000	1.000	1.000	1.000
Leafhoppers	Sticky traps (i a)	−30%	0.501	0.907	0.987	0.999	1.000
Leafhoppers	Sticky traps (i a)	−20%	0.237	0.536	0.743	0.866	0.934
Aphids	Sticky traps (i a)	−50%	0.484	0.894	0.984	0.998	1.000
Aphids	Sticky traps (i a)	−30%	0.174	0.386	0.565	0.703	0.804
Aphids	Sticky traps (i a)	−20%	0.098	0.182	0.265	0.346	0.423
Planthoppers	Sticky traps (i a)	−50%	0.453	0.868	0.975	0.996	0.999
Planthoppers	Sticky traps (i a)	−30%	0.165	0.364	0.537	0.674	0.777
Planthoppers	Sticky traps (i a)	−20%	0.095	0.172	0.250	0.326	0.398
Thrips	Sticky traps	−50%	0.759	0.994	1.000	1.000	1.000
Thrips	Sticky traps	−30%	0.297	0.655	0.850	0.940	0.977
Thrips	Sticky traps	−20%	0.151	0.326	0.483	0.615	0.719
Corn rootworms	Sticky traps	−50%	0.771	0.996	1.000	1.000	1.000
Corn rootworms	Sticky traps	−30%	0.331	0.714	0.896	0.966	0.990
Corn rootworms	Sticky traps	−20%	0.177	0.395	0.577	0.716	0.815
Corn rootworms	Visual counts	−50%	0.595	0.955	0.996	1.000	1.000
Corn rootworms	Visual counts	−30%	0.219	0.493	0.692	0.821	0.899
Corn rootworms	Visual counts	−20%	0.120	0.242	0.358	0.464	0.558
Flea beetles	Visual counts (i a)	−50%	0.258	0.582	0.788	0.900	0.956
Flea beetles	Visual counts (i a)	−30%	0.104	0.200	0.294	0.383	0.467
Flea beetles	Visual counts (i a)	−20%	0.070	0.105	0.140	0.176	0.211
Flea beetles	Sticky traps (i a)	−50%	0.543	0.916	0.985	0.998	1.000
Flea beetles	Sticky traps (i a)	−30%	0.199	0.444	0.629	0.758	0.843
Flea beetles	Sticky traps (i a)	−20%	0.107	0.207	0.304	0.395	0.478
Harvestmen	Pitfall traps	−50%	0.258	0.580	0.786	0.899	0.955
Harvestmen	Pitfall traps	−30%	0.106	0.203	0.299	0.391	0.476
Harvestmen	Pitfall traps	−20%	0.072	0.108	0.146	0.183	0.220
Springtails	Pitfall traps (i a)	−50%	0.781	0.982	0.999	1.000	1.000
Springtails	Pitfall traps (i a)	−30%	0.562	0.937	0.992	0.999	1.000
Springtails	Pitfall traps (i a)	−20%	0.176	0.382	0.541	0.654	0.735
Entomobryids	Pitfall traps (i a)	−50%	0.873	1.000	1.000	1.000	1.000
Entomobryids	Pitfall traps (i a)	−30%	0.379	0.784	0.938	0.984	0.996
Entomobryids	Pitfall traps (i a)	−20%	0.181	0.405	0.590	0.730	0.827
Sap beetles	Pitfall traps	−50%	0.765	0.979	0.999	1.000	1.000
Sap beetles	Pitfall traps	−30%	0.362	0.710	0.854	0.924	0.962
Sap beetles	Pitfall traps	−20%	0.186	0.407	0.571	0.685	0.764
Crickets	Pitfall traps (i a)	−50%	0.649	0.832	0.930	0.973	0.991
Crickets	Pitfall traps (i a)	−30%	0.129	0.267	0.397	0.514	0.615
Crickets	Pitfall traps (i a)	−20%	0.081	0.133	0.186	0.238	0.290
Frit flies	Sticky traps	−50%	0.951	1.000	1.000	1.000	1.000
Frit flies	Sticky traps	−30%	0.502	0.908	0.987	0.998	1.000
Frit flies	Sticky traps	−20%	0.240	0.544	0.750	0.872	0.938
Fungus gnats	Sticky traps	−50%	0.545	0.935	0.993	0.999	1.000
Fungus gnats	Sticky traps	−30%	0.196	0.441	0.635	0.773	0.864
Fungus gnats	Sticky traps	−20%	0.108	0.210	0.310	0.404	0.492
Ants	Pitfall traps	−50%	0.156	0.334	0.479	0.592	0.675
Ants	Pitfall traps	−30%	0.078	0.126	0.174	0.222	0.268
Ants	Pitfall traps	−20%	0.061	0.079	0.098	0.116	0.135
Spiders	Visual counts	−50%	0.359	0.737	0.896	0.959	0.984
Spiders	Visual counts	−30%	0.139	0.292	0.431	0.549	0.645
Spiders	Visual counts	−20%	0.085	0.145	0.206	0.265	0.323
Spiders	Pitfall traps	−50%	0.852	0.999	1.000	1.000	1.000
Spiders	Pitfall traps	−30%	0.352	0.767	0.917	0.976	0.993
Spiders	Pitfall traps	−20%	0.168	0.371	0.546	0.684	0.786
Spiders	Pitfall traps (i)	−50%	0.314	0.687	0.877	0.956	0.985
Spiders	Pitfall traps (i)	−30%	0.122	0.247	0.368	0.478	0.575
Spiders	Pitfall traps (i)	−20%	0.078	0.126	0.174	0.223	0.270
Wolf spiders	Pitfall traps	−50%	0.920	1.000	1.000	1.000	1.000
Wolf spiders	Pitfall traps	−30%	0.445	0.860	0.972	0.996	0.999
Wolf spiders	Pitfall traps	−20%	0.211	0.477	0.678	0.812	0.895
Centipedes	Pitfall traps	−50%	0.227	0.513	0.716	0.843	0.916
Centipedes	Pitfall traps	−30%	0.097	0.179	0.260	0.339	0.413
Centipedes	Pitfall traps	−20%	0.068	0.099	0.129	0.160	0.191
Green lacewings	Visual counts (e)	−50%	0.225	0.510	0.715	0.844	0.918
Green lacewings	Visual counts (e)	−30%	0.096	0.177	0.258	0.336	0.410
Green lacewings	Visual counts (e)	−20%	0.068	0.100	0.131	0.163	0.194
Minute pirate bugs	Sticky traps	−50%	0.922	1.000	1.000	1.000	1.000
Minute pirate bugs	Sticky traps	−30%	0.442	0.857	0.971	0.995	0.999
Minute pirate bugs	Sticky traps	−20%	0.209	0.474	0.674	0.808	0.892

Continued on following page

Appendix 1. Continued

Common name	Sampling method (life stages)[a]	Effect size[b]	Number of replicates				
			3	6	9	12	15
Ground beetles	Pitfall traps	−50%	0.614	0.966	0.998	1.000	1.000
Ground beetles	Pitfall traps	−30%	0.235	0.532	0.738	0.862	0.931
Ground beetles	Pitfall traps	−20%	0.124	0.254	0.378	0.490	0.589
Ground beetles	Pitfall traps (l)	−50%	0.164	0.360	0.529	0.664	0.766
Ground beetles	Pitfall traps (l)	−30%	0.082	0.136	0.191	0.246	0.299
Ground beetles	Pitfall traps (l)	−20%	0.063	0.085	0.106	0.127	0.149
Rove beetles	Pitfall traps	−50%	0.323	0.702	0.887	0.962	0.988
Rove beetles	Pitfall traps	−30%	0.123	0.251	0.374	0.485	0.584
Rove beetles	Pitfall traps	−20%	0.078	0.126	0.174	0.222	0.270
Ladybird beetles	Visual counts	−50%	0.908	1.000	1.000	1.000	1.000
Ladybird beetles	Visual counts	−30%	0.435	0.848	0.968	0.994	0.999
Ladybird beetles	Visual counts	−20%	0.209	0.493	0.670	0.804	0.887
Ladybird beetles	Visual counts (l)	−50%	0.593	0.958	0.997	1.000	1.000
Ladybird beetles	Visual counts (l)	−30%	0.206	0.466	0.665	0.801	0.886
Ladybird beetles	Visual counts (l)	−20%	0.109	0.212	0.313	0.409	0.497
Ladybird beetles	Sticky traps	−50%	0.744	0.989	1.000	1.000	1.000
Ladybird beetles	Sticky traps	−30%	0.312	0.670	0.853	0.936	0.972
Ladybird beetles	Sticky traps	−20%	0.158	0.343	0.503	0.631	0.730
Long-legged flies	Sticky traps	−50%	0.487	0.883	0.976	0.995	0.999
Long-legged flies	Sticky traps	−30%	0.183	0.407	0.588	0.722	0.815
Long-legged flies	Sticky traps	−20%	0.103	0.197	0.287	0.374	0.455
Fairyflies	Sticky traps	−50%	0.677	0.983	0.999	1.000	1.000
Fairyflies	Sticky traps	−30%	0.269	0.603	0.807	0.914	0.964
Fairyflies	Sticky traps	−20%	0.140	0.296	0.441	0.566	0.670
Scelionids	Sticky traps	−50%	0.744	0.993	1.000	1.000	1.000
Scelionids	Sticky traps	−30%	0.287	0.638	0.838	0.933	0.974
Scelionids	Sticky traps	−20%	0.143	0.306	0.455	0.583	0.688
Braconids	Sticky traps	−50%	0.368	0.741	0.892	0.954	0.981
Braconids	Sticky traps	−30%	0.136	0.284	0.420	0.537	0.633
Braconids	Sticky traps	−20%	0.083	0.139	0.196	0.252	0.306
Trichogrammatids	Sticky traps	−50%	0.725	0.980	0.999	1.000	1.000
Trichogrammatids	Sticky traps	−30%	0.295	0.635	0.816	0.907	0.953
Trichogrammatids	Sticky traps	−20%	0.149	0.318	0.466	0.588	0.684

[a] Adults unless otherwise indicated. Life stages (eggs, immatures, adults) combined if inside parentheses.
[b] Percentage change in abundance relative to an experimental control.